

Dataverse

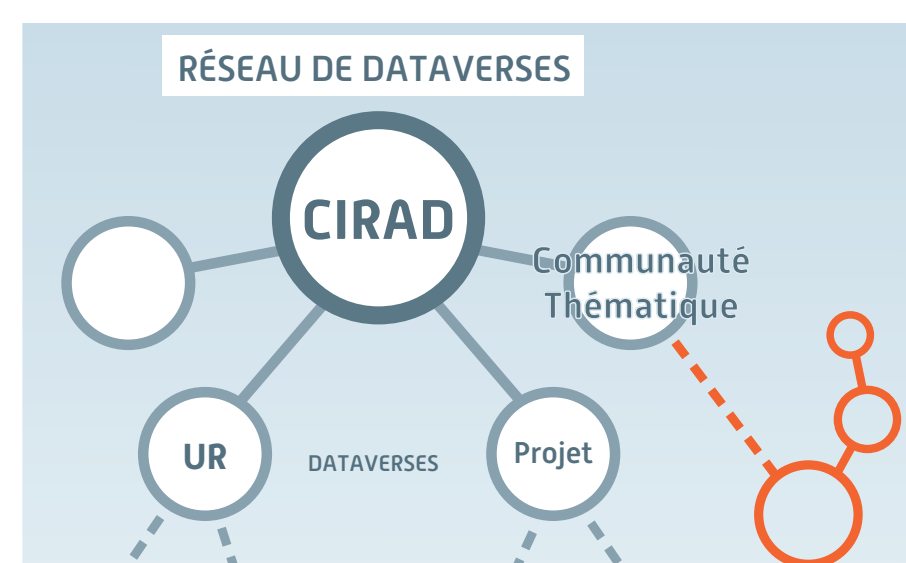
Une plateforme collaborative *open source* pour déposer, décrire, partager, préserver, citer, faire découvrir et réutiliser des données de recherche.

- Appuie les étapes du **cycle de vie** des données¹
- Utilisée par **170 universités et institutions** et **130 revues scientifiques** dans le monde
- Développée par l'**Université d'Harvard**² depuis 2006.

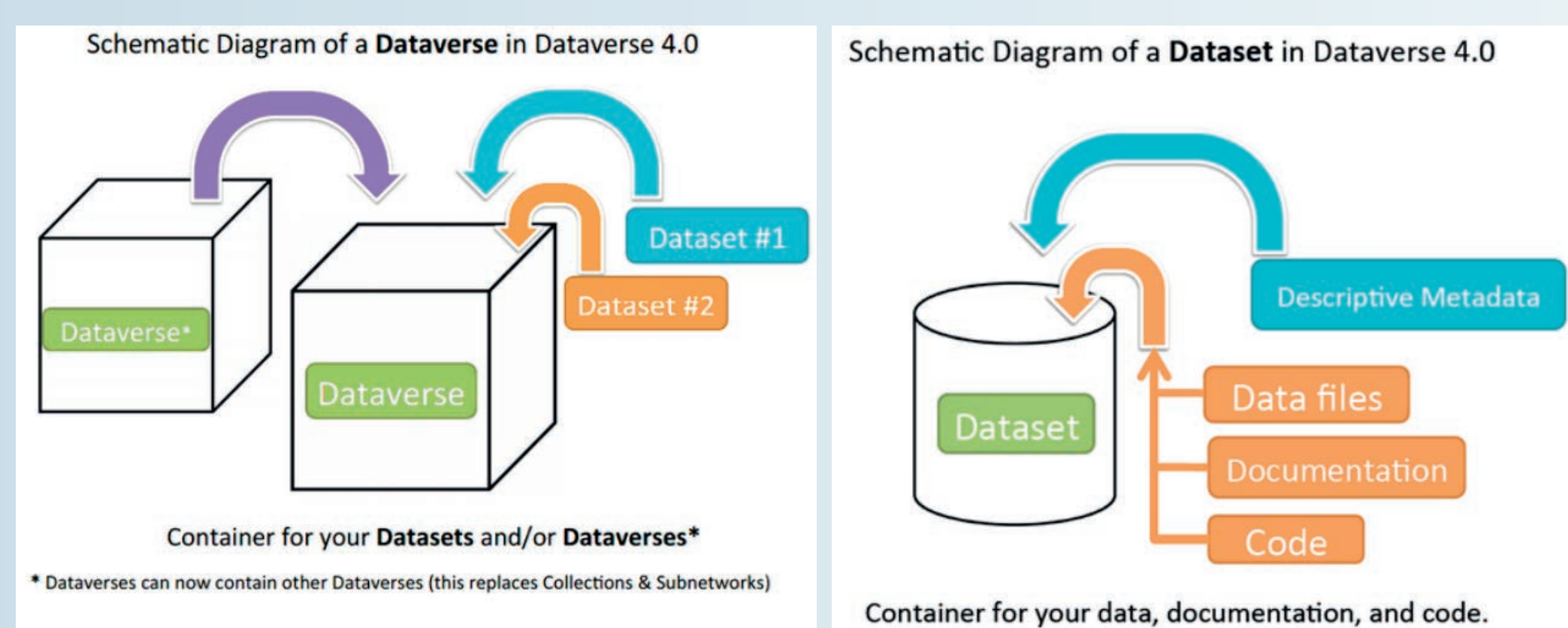
À chaque collectif son dataverse avec ses règles de gestion

Institution, unité de recherche, collectif projet, communauté thématique ou disciplinaire.

- Dataverse permet de se doter de règles de gouvernance à multi-niveaux, pour s'organiser autour des données
- Dispose d'*espaces privés*
- Facilite la mise en réseau des collectifs et rend visible leurs données - espace public -.



Un **Dataverse** est un *container* qui héberge des **Datasets** et d'autres **Dataverses**

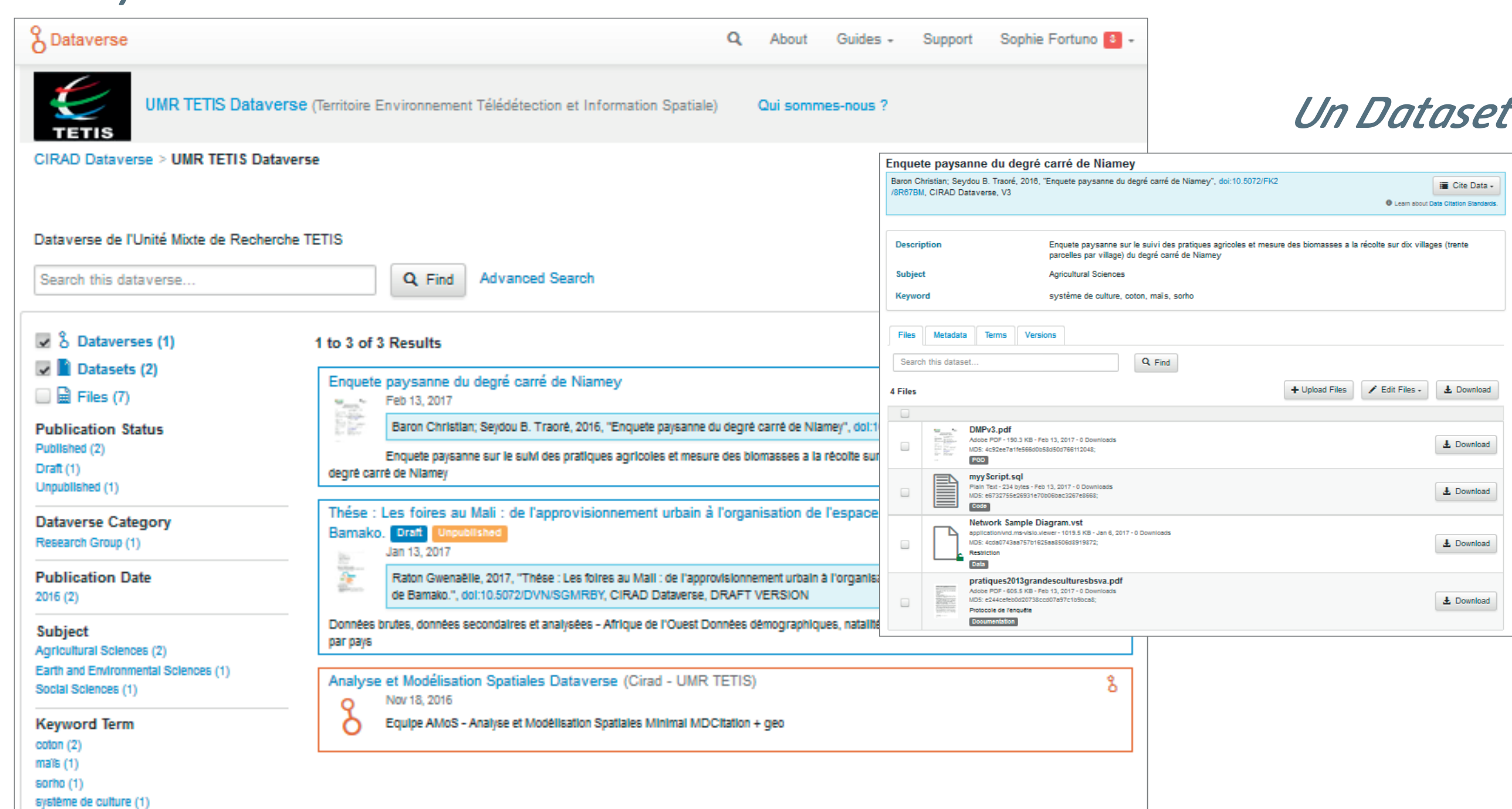


Un **Dataset** est décrit par des **métadonnées**, il reçoit :

- Des fichiers de **données**, du **code** et de la **documentation** associée

Source : <http://dataverse.org>

Exemple de Dataverse



Source : <http://dataverse.cirad.fr>



Pour qui ?

- Chercheurs
- Gestionnaires de données
- Manager d'UR

Dataverse : entrepôt institutionnel sécurisé pour les jeux de données

Publier, diffuser, réutiliser les données

- Fixer les **conditions d'utilisation** de ses données
- **Mettre à disposition** les données d'une **publication** pour une **revue scientifique** ou un **bailleur**
- **Découvrir** les données de partenaires, **consulter** leurs **conditions d'utilisation** et les **télécharger**.

Rendre visible, créditer les données

Pour que le jeu de données soit cité correctement

- Dataverse crée un **identifiant pérenne** du jeu de données (**DOI : Digital Object Identifier**)
- Chaque auteur est identifié avec son affiliation
- Dataverse crée la **citation du jeu** selon un standard international : *auteurs, année, titre du jeu, nom de l'entrepôt, version, identifiant pérenne*.

Viser l'interopérabilité

- Choisir ses **métadonnées** selon une norme ou un standard international adapté à son domaine, sa communauté. Exemples : *Data Documentation Initiative (DDI)*, format international *Dublin Core*
- Déposer des fichiers de données, du code et de la documentation dans un **Dataset** du **Dataverse**.

Exemples fichiers :



Entreposer, préserver les données

- **Pendant le projet**, déposer et partager entre partenaires les fichiers de données de manière **sécurisée** (*compte Cirad ou création de compte*).
- **Après le projet**, maintenir les fichiers : données, code et documentation associée pour une réutilisation.

FORTUNO Sophie¹, JUANES Xavier², GAILLARD Cédric³, ROUVIERE Xavier⁴, FILY Marie-Françoise⁵, BENEDET Fabrice⁶, PARADIS Sébastien⁷

Ces auteurs sont membres du groupe de travail *Patrimoine Numérique Scientifique*

CIRAD : ¹UMR Tetis, ²UMR Selmet, ³UMR Moisa, ⁴Dsi, ⁵Dist, ⁶UR Forêts et Sociétés, ⁷UR BioWooEB

Références

¹Cirad, 2016. Intégrer la gestion des données aux activités de recherche : poster. <http://coop-ist.cirad.fr/actualites/integrer-la-gestion-des-donnees-aux-activites-de-recherche-poster>

² Dataverse Harvard : <http://dataverse.harvard.edu/>